

Interplay between Identification and Optimization in Run-to-run Optimization Schemes

B. Srinivasan and D. Bonvin

Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland.

Abstract

The use of measurements to compensate for model uncertainty and disturbances has received increasing attention in the context of process optimization. The standard procedure consists of iteratively using the measurements for identifying the model parameters and the updated model for optimization. However, in the presence of model mismatch, this scheme suffers from lack of synergy between the identification and optimization problems.

This paper investigates the performance of run-to-run optimization schemes and proposes to modify the objective function of the *identification* problem so as to include the cost function and the constraints of the *optimization* problem. The weights of the various terms in the extended objective function are based on Lagrange multipliers. The performance improvement obtained with the proposed methodology is illustrated *via* the simulation of a semi-batch reaction system.

Keywords: Optimization, Parameter identification, Run-to-run Optimization, Batch Reaction Systems.

1 Introduction

Optimization has received increasing attention in industry since it provides a unified framework to improve productivity without violating safety and quality constraints. However, since most optimization techniques are model-based, and reliable models are rarely available at the industrial level, modeling errors are invariably present. The other major source of uncertainty is the presence of disturbances that arise naturally due for example to variations in initial conditions and process operation. Thus, optimal operation under uncertainty requires either a robust optimization approach or the use of measurements to adjust the off-line calculated strategy [21].

Measurement-based optimization approaches can be classified into indirect and direct schemes, depending on whether or not a model of the system is used for adaptation [4]. In indirect (or model-based) schemes,

a model of the process is updated using measurements and used to compute the optimal inputs [18]. In contrast, direct schemes use the measurements to update the inputs directly, i.e., without the help of a model [5]. Unfortunately, each scheme has a major drawback: Indirect schemes typically suffer from model mismatch, while direct ones often exhibit slow convergence.

The basic philosophy of iterating between identification and optimization can be used towards the on-line optimization of time-varying continuous and batch processes (the performance update is along the time t) or towards the run-to-run optimization of batch processes (the performance update is along the run index k). There are numerous studies in the literature that document the approach, for example in the context of on-line optimization and model-predictive control ([7], [19], [12], [10], [14] and [17]) or with respect to run-to-run optimization ([8], [9], [13], [11], [6] and [20]).

Run-to-run optimization schemes adapt the input profiles of the current run based on measurements from previous ones. These schemes typically exploit the fact that many industrial processes, especially in the fields of batch chemical production [4] and semiconductor manufacturing [16], are repeated over time.

This paper considers only *model-based run-to-run* optimization schemes that involve the two steps of identification and optimization. These schemes work well provided: (i) the excitation generated by the optimal inputs is sufficient to identify all uncertain parameters, and (ii) there is no model mismatch. Since these two assumptions are difficult to verify in practice, there is often lack of synergy between the identification and the optimization steps ([18], [1]), and a tradeoff between exploitation (optimization) and exploration (identification) is needed [15].

The situation is somewhat similar to that found in the area of system identification and control. The tasks of identification and control, which are typically antinomic (dual control problem [22]), should not be solved independently but rather hand in hand. For best performance, matching-criteria schemes have been proposed (identification for control, see for example [2]).

Along the same lines, this paper proposes matching objective functions for the identification and optimization problems.

The underlying idea of this work is “modeling for optimization”. Usually, the goal of identification is to come up with a model that, for given inputs, can predict the outputs of the reality. In contrast, the goal here is to identify a model that allows computing inputs that are nearly optimal for the reality. For this, the cost and constraints of the optimization problem will be included in the identification objective. The necessary weighting between the different terms is based on the Lagrange multipliers of the optimization problem.

The paper is organized as follows. Section 2 introduces the standard model-based run-to-run optimization scheme and illustrates the absence of synergy between the identification and optimization tasks. A modification of the identification objective is proposed in Section 3. Section 4 evaluates the performance of the resulting scheme via simulation of a semi-batch reaction system, and conclusions are drawn in Section 5.

2 Standard Model-based Run-to-run Optimization Scheme

In the context of model-based run-to-run optimization, it is important to distinguish between the reality and the model. Let the *reality* be governed by the following equations:

$$\dot{x} = F(x, \theta, u) + d(t), \quad x(0) = x_0, \quad y = h(x, \theta) + v(t) \quad (1)$$

where x are the states, x_0 the initial conditions, u the inputs, y the measured outputs, θ the parameters, $F(x, \theta, u)$ the system equations, $h(x, \theta)$ the output equations, and $d(t)$ and $v(t)$ the zero-mean process and measurement noises, respectively.

A *model* of the reality is used to carry out the optimization. The accent $\bar{\cdot}$ will denote quantities related to the model:

$$\dot{\bar{x}} = \bar{F}(\bar{x}, \bar{\theta}, u), \quad \bar{x}(0) = x_0, \quad \bar{y} = \bar{h}(\bar{x}, \bar{\theta}) \quad (2)$$

Note that the inputs to the reality and to the model are the same, while the states and outputs are different. This, on the one hand, is due to not using the unknown d and v in (2). On the other hand, the parameters might be inaccurate, and the system and output equations could also be different, $F \neq \bar{F}$ and $h \neq \bar{h}$. The latter is in fact true when there is model mismatch.

Model-based run-to-run optimization scheme involves the following two steps:

- *Identification:* Identify the parameters of a given model structure using measurements from the previous run. Typically, the following least-squares minimization problem is solved to compute the optimal parameters after the k^{th} run:

$$\begin{aligned} \bar{\theta}_k^* &= \arg \min_{\bar{\theta}} \bar{J}_k^{\text{iden}} \\ \bar{J}_k^{\text{iden}} &= \sum_{i=1}^{N_k} \|y_k(i) - \bar{y}_k(i)\| \\ \text{s.t.} \quad &\dot{\bar{x}}_k = \bar{F}(\bar{x}_k, \bar{\theta}, u_k^*), \quad \bar{x}_k(0) = x_{k_0} \\ &\bar{y}_k = \bar{h}(\bar{x}_k, \bar{\theta}), \quad \text{given } y_k(i), \forall i \leq N_k \end{aligned} \quad (3)$$

where the superscript $*$ indicates a quantity calculated via optimization, the subscript k a quantity associated with the k^{th} run, for example, u_k^* the optimal input, and N_k the number of measurements in the k^{th} run.

- *Optimization:* Numerically compute the optimal input profiles for the $(k+1)^{\text{st}}$ run on the basis of the model estimated after the k^{th} run:

$$\begin{aligned} u_{k+1}^* &= \arg \min_u \bar{J}_{k+1}^{\text{opt}} \\ \bar{J}_{k+1}^{\text{opt}} &= \phi(\bar{x}_{k+1}(t_f), \bar{\theta}_k^*) \\ \text{s.t.} \quad &\dot{\bar{x}}_{k+1} = \bar{F}(\bar{x}_{k+1}, \bar{\theta}_k^*, u), \quad \bar{x}_{k+1}(0) = x_{k+1_0} \\ &S(\bar{x}_{k+1}, \bar{\theta}_k^*, u) \leq 0, \quad T(\bar{x}_{k+1}(t_f), \bar{\theta}_k^*) \leq 0 \end{aligned} \quad (4)$$

where ϕ is the terminal cost, t_f the final time, and S and T the path and terminal constraints, respectively. For notational simplicity, define $\bar{\phi}_{k+1} \equiv \phi(\bar{x}_{k+1}(t_f), \bar{\theta}_k^*)$, $\bar{S}_{k+1} \equiv S(\bar{x}_{k+1}, \bar{\theta}_k^*, u_{k+1}^*)$, and $\bar{T}_{k+1} \equiv T(\bar{x}_{k+1}(t_f), \bar{\theta}_k^*)$.

The iterative procedure for model-based run-to-run adaptation starts with $k = 0$ and some input profiles u_0^* . Then, each iteration solves the identification problem (3) for the parameters $\bar{\theta}_k^*$, plugs their value in the optimization problem (4) to obtain the inputs u_{k+1}^* and then sets $k := k + 1$. This procedure is repeated until $\bar{J}_{k+1}^{\text{opt}}$ has converged.

Persistency of excitation is an important concept in the context of parameter identification [3]. Mathematically, this corresponds to the hessian $\frac{\partial^2 \bar{J}_k^{\text{iden}}}{\partial \bar{\theta}^2}$ being positive definite so that (3) has a unique solution. This hessian depends on the inputs that are applied to the system. A set of inputs is persistently exciting if the hessian is indeed positive definite. In other words, persistently exciting inputs show enough variability to uncover the model parameters that have to be identified.

If u_0^* is persistently exciting, then, in the absence of model mismatch and noise, the parameters computed by solving (3) are in fact the true parameters ($\bar{\theta}_0^* = \theta_{\text{true}}$). Furthermore, the inputs computed by solving (4) are optimal ($u_1^* = u_{\text{optimal}}$). Thus, the

scheme converges in a single iteration. In the presence of noise, it is not $\bar{\theta}_0^*$ but its average value over different realizations that converges to θ_{true} . In the presence of model mismatch, θ_{true} does not exist, and the parameters converge to values that minimize errors in output prediction. However, there is no guarantee that the parameters that give the best output prediction will also appropriately predict the cost and constraints of the optimization problem. And, if the cost and constraints of the optimization problem are not predicted properly, the solution obtained by optimizing the *model* in (4) does not optimize the *reality*!

In summary, the model-based run-to-run optimization schemes necessitate two core assumptions for proper functioning:

1. *Persistent excitation of the optimal inputs* - Yet, there is no guarantee that the optimal profiles are indeed persistently exciting. This problem becomes more acute with a large number of uncertain parameters.
2. *Absence of model mismatch* - Yet, some of the dynamics are invariably left out in the modeling process. In addition, parameters are often inaccurate.

These two problems are in fact linked since, when the excitation is insufficient, some parameters have to be kept at their nominal values, thus leading to model mismatch. In the sequel, it will be assumed that only as many parameters as allowed by the persistency of excitation of the inputs are identified. Thus, the problem of model mismatch caused either by unmodeled dynamics, or by fixing some of the parameters at inaccurate values, will be addressed.

3 Objective Function for the Identification Problem

There is a difference between optimizing the *reality* (1) and optimizing the *model* (2). The problem (4) optimizes the model under the assumption that the model is ‘close’ to the reality (certainty-equivalence principle [3]). However, this assumption is not valid when there is model mismatch, for which case it is important to choose a model that predicts well not only the state evolution, but also the cost and the constraints of the optimization problem.

3.1 Necessary Conditions of Optimality

Consider the problem of maximizing the cost $\phi(x(t_f), \theta)$ of the reality subject to constraints on the real state evolution. The problem formulation is similar to (4), but expressed for the reality instead of the model. The adjointed cost for the reality is given by:

$$\begin{aligned} \mathcal{J}_k^{opt} &= \phi(x_k(t_f), \theta) + \nu_k^T T(x_k(t_f), \theta) \\ &+ \int_0^{t_f} \mu_k^T S(x_k, \theta, u_k) dt \\ &+ \int_0^{t_f} \lambda_k^T (F(x_k, \theta, u_k) + d - \dot{x}_k) dt \end{aligned} \quad (5)$$

where $\lambda_k(t)$, $\mu_k(t)$, and ν_k are the Lagrange multipliers associated with the system equations, the path and terminal constraints, respectively. Note that μ and ν are positive if S and T are zero and zero otherwise. So, the terms $\mu^T S$ and $\nu^T T$ are always zero when there is no constraint violation. Furthermore, the last term is always zero since, by definition, the evolution of the states satisfies the differential equations. However, this last term is kept since it will help relate the errors in state prediction to the optimal cost, as discussed later.

The necessary conditions of optimality are given by $\frac{\partial \mathcal{J}_k^{opt}}{\partial u} = 0$, which in turn can be expressed as the sum of two terms:

$$\frac{\partial \mathcal{J}_k^{opt}}{\partial u} = \frac{\partial \bar{\mathcal{J}}_k^{opt}}{\partial u} + \frac{\partial (\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt})}{\partial u} = 0 \quad (6)$$

where $\bar{\mathcal{J}}_k^{opt}$ is the adjointed cost for the model.

Optimizing the reality is equivalent to satisfying (6). Among the two terms therein, the optimization problem (4) forces the first term $\frac{\partial \bar{\mathcal{J}}_k^{opt}}{\partial u}$ to zero, while the identification problem (3) should be reformulated in such a way that it pushes the second term $\frac{\partial (\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt})}{\partial u}$ to zero.

3.2 Modification of the Identification Problem

Since the gradient term $\frac{\partial (\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt})}{\partial u}$ is difficult to construct from the measurements, $|\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt}|$ is proposed as the objective function for the identification problem. Note that the proposed objective function is only a heuristic substitute for the desired one, with the rationale being: Once $|\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt}|$ is pushed to zero for all u in the neighborhood of the current operating point, then its derivatives with respect to u will also be close to zero.

$\bar{\mathcal{J}}_k^{opt}$ can be computed directly from the model. The evaluation of \mathcal{J}_k^{opt} given by (5) is more difficult and is discussed next. First, the actual measurements for $\phi_k \equiv \phi(x_k(t_f), \theta)$, $T_k \equiv T(x_k(t_f), \theta)$ and $S_k \equiv S(x_k, \theta, u_k)$ are necessary. These are typically available since ϕ_k corresponds to the desired cost at final time, and the path and terminal constraints are routinely measured for security and quality reasons. The Lagrange multipliers can be approximated using the model, i.e., $\lambda_k \approx \bar{\lambda}_k$, $\mu_k \approx \bar{\mu}_k$, and $\nu_k \approx \bar{\nu}_k$, where:

$$\bar{\lambda}_k^T = -\bar{\lambda}_k^T \frac{\partial F}{\partial x} + \bar{\mu}_k^T \frac{\partial S}{\partial x}, \quad \bar{\lambda}_k^T(t_f) = \left(\frac{\partial \phi}{\partial x} + \bar{\nu}_k^T \frac{\partial T}{\partial x} \right) \Big|_{t_f} \quad (7)$$

Then,

$$\begin{aligned} \mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt} &= \phi_k - \bar{\phi}_k + \bar{\nu}_k^T (T_k - \bar{T}_k) \\ &+ \int_0^{t_f} \bar{\mu}_k^T (S_k - \bar{S}_k) dt \\ &+ \int_0^{t_f} \bar{\lambda}_k^T (F + d - \dot{x}_k - \bar{F} + \dot{\bar{x}}_k) dt \end{aligned} \quad (8)$$

Under the assumption that $F + d \approx \bar{F}$, the last term can be made dependent upon the difference in the states (rather than their derivatives) using the product rule of integration,

$$\begin{aligned} \int_0^{t_f} \lambda^T (\dot{\bar{x}}_k - \dot{x}_k) dt &= \int_0^{t_f} \dot{\lambda}^T (x_k - \bar{x}_k) dt \\ &- \lambda(t_f)^T (x_k(t_f) - \bar{x}_k(t_f)) + \lambda(0)^T (x_k(0) - \bar{x}_k(0)) \end{aligned} \quad (9)$$

Next, since the states of the reality are not measured, a least-squares estimate of $(x_k - \bar{x}_k)$ is obtained: $(x_k - \bar{x}_k) = \left(\frac{\partial h}{\partial x}\right)^+ (y_k - \bar{y}_k)$, where the superscript $+$ denotes the pseudo inverse. Using the triangle inequality:

$$\begin{aligned} |\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt}| &\leq |\phi_k - \bar{\phi}_k| + |\bar{\nu}_k^T (T_k - \bar{T}_k)| + \\ &\sum_{i=1}^{N_k} |\bar{\mu}_k^T(i) (S_k(i) - \bar{S}_k(i))| + \\ &|\bar{\lambda}_k^T(N_k) \left(\frac{\partial h}{\partial x}\right)^+ (y_k(N_k) - \bar{y}_k(N_k))| + \\ &\sum_{i=1}^{N_k} |\bar{\lambda}_k^T(i) \left(\frac{\partial h}{\partial x}\right)^+ (y_k(i) - \bar{y}_k(i))| \end{aligned} \quad (10)$$

The upper bound for $|\mathcal{J}_k^{opt} - \bar{\mathcal{J}}_k^{opt}|$ in (10) contains various terms that express the difference between the reality and the model: (A) The last two terms represent the accuracy of the system states, and (B) the other terms are associated with the accuracy of the cost and the constraints of the optimization problem. It can be seen that (3) has only type A terms, and to have synergy between the identification and optimization problems, additional type B terms need to be introduced. Thus, the proposed identification problem reads:

$$\begin{aligned} \bar{\theta}_k^* &= \arg \min_{\bar{\theta}} \bar{J}_k^{iden} \\ \bar{J}_k^{iden} &= |\phi_k - \bar{\phi}_k| + |\alpha^T (T_k - \bar{T}_k)| + \\ &\sum_{i=1}^{N_k} |\beta^T (S_k(i) - \bar{S}_k(i))| + \sum_{i=1}^{N_k} |\gamma^T (y_k(i) - \bar{y}_k(i))| \\ \text{s.t. } \dot{\bar{x}}_k &= \bar{F}(\bar{x}_k, \bar{\theta}, u_k^*), \quad \bar{x}_k(0) = x_{k_0} \\ \bar{y}_k &= \bar{h}(\bar{x}_k, \bar{\theta}), \quad \text{given } y_k(i), \forall i \leq N_k \end{aligned} \quad (11)$$

where α , β , and γ are weighting parameters.

Note that, in the absence of model mismatch and noise, the optimization gives $\bar{\theta}_0^* = \theta_{true}$ independent of the weights that are used. This results from $\phi_i = \bar{\phi}_i$, $y_i =$

\bar{y}_i , $T_i = \bar{T}_i$, and $S_i = \bar{S}_i$, i.e., all the terms are zero. However, in the presence of model mismatch, $\bar{\theta}_k^*$ will depend on the relative weights between the different terms.

Two extreme cases of weighting are discussed next:

1. *Having only terms of type A* - The state prediction is quite good, but the prediction of the cost and of the constraints of the optimization problem can be poor.
2. *Having only terms of type B* - Though the cost and constraints of the optimization problem for the given inputs are well predicted, there is no guarantee that the prediction will be good for variations in the inputs. In other words, the extrapolation capability of the resulting model could be quite poor.

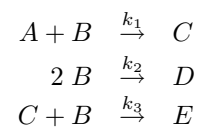
Thus, it is important to strike a balance between the two types of terms in the objective function of the identification problem (11) for a good synergy between the identification and optimization tasks. The best weighting between the different terms is based on the Lagrange multipliers of the optimization problem as illustrated in (10), i.e., $\alpha = \nu_k$, $\beta = \mu_k$, and $\gamma = \dot{\lambda}_k$. Alternatively, these weights could be considered as tuning parameters that are adjusted to obtain the desired performance.

Except for the proposed modification of the identification objective, the iterative procedure remains exactly the same. It consists of iteratively solving the modified identification problem (11) and the optimization problem (4).

4 Run-to-run Optimization of a Semi-batch Reactor

The performance of the standard and proposed run-to-run optimization schemes will be compared in simulation. The isothermal semi-batch reaction system for the acetoacetylation of pyrrole with diketene is considered. Pyrrole is initially in the reactor and diketene is fed so as to maximize the production of acetoacetyl pyrrole. The reaction system, the model and the optimization are briefly presented next. For a more detailed description, the reader is referred to [19].

The simulated reality has the three reactions:



where A: pyrrole, B: diketene, C: 2-acetoacetyl pyrrole, D: dehydroacetic acid, E: undesired by-product. The

k_1	0.1	l/mol min
k_2	0.2	l/mol min
k_3	0.03	l/mol min
$c_{b_{in}}$	5	mol/l
$c_{unf,max}$	0.15	mol/l
u_{min}	0	l/min
u_{max}	0.002	l/min
t_f	120	min
c_{a_0}	1	mol/l
c_{b_0}	0	mol/l
V_0	1	l

Table 1: Parameter values and initial conditions

optimization problem reads:

$$\begin{aligned}
\max_{u(t)} J &= c_c(t_f)V(t_f) \\
s.t. \quad &\dot{c}_a = -k_1 c_a c_b - (u/V) c_a \\
&\dot{c}_b = -k_1 c_a c_b - 2 k_2 c_b^2 - k_3 c_b c_c \\
&\quad + (u/V) (c_{b_{in}} - c_b) \\
&\dot{c}_c = k_1 c_a c_b - k_3 c_b c_c - (u/V) c_c \\
&\dot{c}_d = k_2 c_b^2 - (u/V) c_d \\
&\dot{c}_e = k_3 c_a c_b - (u/V) c_e \\
&\dot{V} = u \\
&u_{min} \leq u \leq u_{max} \\
&c_d(t_f) + c_e(t_f) \leq c_{unf,max}
\end{aligned} \tag{12}$$

where c_a , c_b , c_c , c_d , and c_e are the concentrations of A , B , C , D , and E in mol/l, respectively. The feed consists of Species B with concentration $c_{b_{in}}$. The goal is to maximize the number of moles of C at the given final time t_f by manipulating the feedrate u (l/min), whilst satisfying bounds on the input and a terminal constraint on the concentrations of the undesired species D and E . The numerical values of the parameters used in the simulation are given in Table 1.

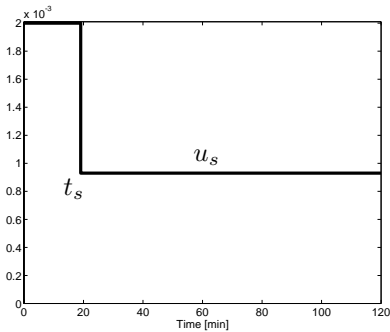


Figure 1: Optimal input (the singular arc is approximated by the constant value u_s)

The optimal solution can be computed numerically (Figure 1). The input is characterized as having two intervals: (i) input at its upper bound u_{max} , and (ii)

Parameter	$k_3^{model} = k_3$		$k_3^{model} = 0$	
Noise	0%	5%	0%	5%
Standard	0.425	0.403	0.357	0.334
Proposed	0.425	0.414	0.424	0.392

Table 2: Adjoined cost [mol] for standard and proposed schemes with/without model mismatch, and with/without measurement noise. The results with noise are averaged over 50 realizations.

input being singular, approximated here by a constant value u_s . The singular arc represents the compromise between producing the desired product C and the undesired side products D and E . From the optimal solution, the natural parameterization corresponds to the following set: t_s , switching time between the input at the upper bound and the singular interval; and u_s , the constant feedrate during the singular interval. The maximum amount of C obtained with this profile is 0.4250 mol, with $t_s = 19.07$ min and $u_s = 9.3 \times 10^{-4}$ l/min. The optimal solution without approximation of the singular interval is $J^{opt} = 0.4258$, and the negligible loss in performance (0.1%) justifies the approximation.

It is assumed that the concentrations of B and C are measured every 20 min. Only the parameters k_1 and k_2 are adapted from run-to-run, using the measurements from only the previous run, while k_3 is kept constant. Two cases of model mismatch are considered: i) no model mismatch $k_3^{model} = k_3$, and ii) $k_3^{model} = 0$, where the third reaction is assumed to be non-existent. Similarly, two levels of noise are considered: a) no noise, and b) 5% multiplicative gaussian measurement noise.

Two schemes are compared: The standard one that consists of iterating between (3) and (4), and the proposed one that consists of iterating between (11) and (4). For the given example, both schemes converge in one or two iterations. Also, in the presence of noise, the solution comes close to its average value in a few iterations.

The values of the adjoined cost for various schemes are presented in Table 2. If the terminal constraint is satisfied, the adjoined cost is the actual number of moles of C . However, if the constraint is violated, the adjoined cost is the number of moles of C discounted (based on the Lagrange multiplier) by the amount of constraint violation.

In the absence of noise and model mismatch (Column 1), there is no difference between the standard and the proposed schemes. The proposed scheme performs marginally better in the presence of noise and no model mismatch (Column 2). However, in the presence of model mismatch, the proposed scheme provides around 20% improvement in cost, both without and with noise (Columns 3 and 4).

It is interesting to compare the values of the parameters k_1 and k_2 to which the two schemes converge in the case of model mismatch and no noise (Column 3). The standard procedure converges to $(k_1, k_2) = (0.105, 0.569)$, while with the modified objective function $(k_1, k_2) = (0.087, 0.454)$ is obtained. Both schemes increase the kinetic coefficient of the second reaction to account for the loss of B being consumed by the third (unmodeled) reaction. However, it is interesting to see that the standard scheme predicts k_1 fairly accurately in order to get a good fit for C , while the proposed scheme underestimates k_1 to account for the reduction in C (the cost) caused by the third reaction.

5 Conclusions

Model-based methods for optimization using measurements in the presence of uncertainty are attractive due to their fast convergence. However, the convergence can be to a point far away from the optimum when there is model mismatch. To handle this problem, a modification of the objective function of the identification problem has been proposed in this paper. The cost and the constraints of the optimization problem are included with appropriate weighting in the identification objective. This way, the reality can be optimized despite the availability of a faulty or incomplete model.

References

- [1] M. Agarwal. Feasibility of on-line reoptimization in batch processes. *Chem. Eng. Comm.*, 158:19–29, 1997.
- [2] K. J. Aström. Matching criteria for control and identification. In *European Control Conference*, pages 248–251, Groningen, Netherlands, 1993.
- [3] K. J. Aström and B. Wittenmark. *Adaptive Control*. Addison-Wesley, Reading, Massachusetts, 2nd edition, 1995.
- [4] D. Bonvin, B. Srinivasan, and D. Ruppen. Dynamic optimization in the batch chemical industry. In *Chemical Process Control - 6*, Tucson, AZ, 2001.
- [5] G. E. P. Box and N. R. Draper. *Empirical Model-building and Response Surfaces*. John Wiley, New York, 1987.
- [6] T. L. Clarke-Pringle and J. F. MacGregor. Optimization of molecular weight distribution using batch-to-batch adjustments. *Ind. Eng. Chem. Res.*, 37:3660–3669, 1998.
- [7] J. W. Eaton and J. B. Rawlings. Feedback control of nonlinear processes using on-line optimization techniques. *Comp. Chem. Eng.*, 14:469–479, 1990.
- [8] C. Filippi-Bossy, J. Bordet, J. Villiermaux, S. Marchal-Brassely, and C. Georgakis. Batch reactor optimization by use of tendency models. *Comp. Chem. Eng.*, 13:35–47, 1989.
- [9] J. Fotopoulos, C. Georgakis, and H. G. Stenger. Uncertainty issues in the modeling and optimisation of batch reactors with tendency modeling. *Chem. Engng. Sci.*, 49:5533–5548, 1994.
- [10] G. Gattu and E. Zafiriou. A methodology for on-line setpoint modification for batch reactor control in the presence of modeling error. *Chem. Eng. Journal*, 75(1):21–29, 1999.
- [11] M. Ge, Q. G. Wang, M. S. Chin, T. H. Lee, C. C. Hang, and K. H. Teo. An effective technique for batch process optimization with application to crystallization. *Trans. IChemE*, 78A:99–106, 2000.
- [12] A. Helbig, O. Abel, and W. Marquardt. Model predictive control for the on-line optimization of semi-batch reactors. In *American Control Conference*, pages 1695–1699, Philadelphia, PA, 1998.
- [13] M. V. Le Lann, M. Cabassud, and G. Casamatta. Modeling, optimization, and control of batch chemical reactors in fine chemical production. In *IFAC DYCOPS-5*, pages 751–760, Corfu, Greece, 1998.
- [14] J. Lee, K. S. Lee, J. H. Lee, and S. Park. An on-line batch span minimization and quality control strategy for batch and semi-batch processes. In *IFAC ADCHEM'00*, pages 705–712, Pisa, Italy, 2000.
- [15] E. C. Martinez. Batch process modeling for optimization and reinforcement learning. *Comp. Chem. Eng.*, 24:1187–1193, 2000.
- [16] J. Moyne, E. del Castillo, and A. M. Hurwitz. *Run-to-run Control in Semiconductor Manufacturing*. CRC Press, Boca Raton, Florida, 2001.
- [17] M. Noda, T. Chida, S. Hasebe, and I. Hashimoto. On-line optimization system of pilot scale multi-effect batch distillation system. *Comp. Chem. Eng.*, 24:1577–1583, 2000.
- [18] P. D. Roberts and T. W. C. Williams. On an algorithm for combined system optimization and parameter estimation. *Automatica*, 17:199–209, 1981.
- [19] D. Ruppen, D. Bonvin, and D. W. T. Rippin. Implementation of adaptive optimal operation for a semi-batch reaction system. *Comp. Chem. Eng.*, 22:185–189, 1998.
- [20] B. Srinivasan, C. J Primus, D. Bonvin, and N. L. Ricker. Run-to-run optimization via generalized constraint control. *Control Eng. Practice*, 9:911–919, 2001.
- [21] P. Terwiesch, M. Agarwal, and D. W. T. Rippin. Batch unit optimization with imperfect modeling - A survey. *J. Process Contr.*, 4:238–258, 1994.
- [22] B. Wittenmark. Adaptive dual control methods: An overview. In *IFAC Symposium on Adaptive Syst. in Control and Signal Proc.*, pages 67–72, Budapest, 1995.